

Daniel Sheppard,¹ Graeme Henkelman,^{1,*} and O. Anatole von Lilienfeld^{2,†}

¹*Department of Chemistry and Biochemistry and Institute for Computational Engineering and Sciences,
The University of Texas at Austin, Austin, TX 78712-0165*

²*Multiscale Dynamic Materials Modeling Department,
Sandia National Laboratories, Albuquerque, New Mexico 87185-1322, USA*

(Dated: July 8, 2010)

Based on molecular grand canonical ensemble density functional theory, we present a theoretical description of how reaction barriers and enthalpies change as atoms in the system are subjected to alchemical transformations, from one element into another. The change in the energy barrier for the umbrella inversion of ammonia is calculated along an alchemical path in which the molecule is transformed into water, and the change in the enthalpy of protonation for methane is calculated as the molecule is transformed into a neon atom *via* ammonia, water, and hydrogen fluoride. Alchemical derivatives are calculated analytically from the electrostatic potential in the unperturbed system, and compared to numerical derivatives calculated with finite difference interpolation of the pseudopotentials for the atoms being transformed. Good agreement is found between the analytical and numerical derivatives. Alchemical derivatives are also shown to be predictive for integer changes in atomic numbers for oxygen binding to a 79-atom palladium nanoparticle, illustrating their potential use in gradient-based optimization algorithms for the rational design of catalysts.

I. INTRODUCTION

Chemical compound space (CCS) can be defined as the set of all combinations of chemical elements in all geometric isomers. Combinatorial exploration of CCS has been used to produce large virtual chemical databases [1, 2] which would then need to be searched to find the compound with the property of interest [3–5]. More efficient approaches to search CCS include simulated annealing [6], genetic algorithms [7], cluster expansions that can be combined with density functional theory [8], optimizations based on alchemical gradients [9–12], or various optimization methods in the coefficient space of linear combinations of atomic potentials [13–18].

The energetics of chemical reactions that involve geometrical changes are relevant for many important phenomena. In this study, we address the effect of varying stoichiometry on properties which are explicitly geometry dependent, such as activation barriers or reaction enthalpies.

First, we have studied the effect of alchemical transformation for simple molecular model systems, for which we can easily assess accuracy and demonstrate numerical feasibility.

In particular, we choose a subset of CCS including only stable compounds containing 10 protons and electrons, such as CH₄, where a central atom is saturated with hydrogens. Table I shows the alchemical paths and properties which have been investigated. Reaction protonation enthalpies along the four continuous transformations, CH₄ → NH₃ → H₂O → HF → Ne, have been computed, as well as the activation energy for the umbrella flipping of ammonia as it is continuously transformed into a water molecule. Analytical alchemical derivatives along each stoichiometrical transformation were calculated and compared to their finite difference analogue.

Secondly, we examine alchemical derivatives for the binding energy of oxygen to a nanoparticle. It has been shown by Nørskov and coworkers [19] that the binding energy of oxygen to a metal surface can be used to predict its catalytic activity for the oxygen reduction reaction. This reaction is important because it limits the kinetics in fuel cell cathodes. We show how alchemical derivatives can be used to navigate the CCS of nanoparticles to tune the oxygen binding energy.

In these examples, we show how analytic alchemical derivatives, which can be evaluated from single point energy calculations, are consistent with finite difference derivatives calculated from interpolation of the pseudopotentials. We also show how well the analytic derivatives are for predicting changes in binding energies and barriers over integer changes in atomic numbers. This is particularly important for the use of alchemical derivatives in gradient based optimization of material properties.

II. THEORY

A. Taylor expansion in chemical space

Free energy differences can be evaluated between any two distinct thermodynamic states through thermodynamic integration [20, 21], and variants of it, such as free energy perturbation methods [22]. Typically, an order parameter, λ , is used

TABLE I: Investigated properties, reaction enthalpies ΔH and energy barriers E^{act} , and chemical species connected by alchemical paths *via* the order parameter λ . ($0 \leq \lambda \leq 1$)

$\lambda=0$	$\lambda=1$	property
CH ₄	NH ₃	ΔH
NH ₃	H ₂ O	$\Delta H, E^{act}$
H ₂ O	HF	ΔH
HF	Ne	ΔH

*Electronic address: henkelman@mail.utexas.edu

†Electronic address: oavonli@sandia.gov

to drive the system from one state to the other. Integration over the statistical mechanical ensemble averages of the force over λ yields the free energy difference between the two states. In the same way, we use here λ to express a compound's potential energy as a Taylor expansion in CCS around a reference compound, where $\lambda = 0$ and the energy is E^0 . The energy of any other compound, for which $\lambda = 1$, is approximated as

$$E(\lambda = 1) = E^0 + \partial_\lambda E^0 d\lambda + \frac{1}{2} \partial_\lambda^2 E^0 d\lambda^2 + \text{HOT} \quad (1)$$

where $d\lambda = 1$, $\partial_\lambda E$ denotes the derivative of E with respect to λ and HOT are the higher order terms.

Within density functional theory (DFT) [23], a compound is defined by its proton density distribution, $Z(\mathbf{r})$, and the number of electrons N_e which—after the self-consistent field cycle—yield the ground state electron density, ρ . These quantities constitute the extensive particle variables within molecular grand canonical DFT [10]. Since we explicitly focus on compounds with different geometries it is now convenient to include the positions of the nuclei $\{\mathbf{R}_I\}$ in the set of extensive variables, assuming the conventional classical point charge distribution of atomic numbers, $Z(\mathbf{r}) = \sum_I N_I \delta(\mathbf{R}_I - \mathbf{r})$, where N_I is the atomic number of atom I .

Consequently, the first order term from Eq. (1) is

$$\begin{aligned} \partial_\lambda E &= \int d\mathbf{r} \delta_{Z(\mathbf{r})} E \partial_\lambda Z(\mathbf{r}) + \sum_I \partial_{\mathbf{R}_I} E \cdot \partial_\lambda \mathbf{R}_I \\ &\quad + \partial_{N_e} E \partial_\lambda N_e \\ &= \int d\mathbf{r} \mu_n(\mathbf{r}) \partial_\lambda Z(\mathbf{r}) - \sum_I \mathbf{F}_I \cdot \partial_\lambda \mathbf{R}_I \\ &\quad + \mu_e \partial_\lambda N_e. \end{aligned} \quad (2)$$

The nuclear chemical potential, $\delta_{Z(\mathbf{r})} E = \mu_n(\mathbf{r})$, is the derivative of the energy with respect to variation in the nuclear charge distribution. At the position of atom I , $\mu_n(\mathbf{R}_I)$ is the derivative of the energy with respect to the nuclear charge $Z(\mathbf{R}_I)$. It is therefore called the ‘‘alchemical potential.’’ The ionic forces, $\{\mathbf{F}_I\}$, are the negative gradients of the energy with respect to atomic position. The electronic chemical potential μ_e is the derivative of the energy with respect to a change in N_e , i.e. the molecular eigenvalue of the highest occupied molecular Kohn-Sham orbital [24, 25].

Analogously, the second order term from Eq. (1) can be obtained, and contains all the pure and mixed second order derivatives such as: the alchemical hardness $\delta_{Z(\mathbf{r})} \delta_{Z(\mathbf{r}')} E = \delta_{Z(\mathbf{r}')} \mu_n(\mathbf{r}) = \eta_n(\mathbf{r}, \mathbf{r}')$; the electronic hardness, $\partial_{N_e} \partial_{N_e} E = \partial_{N_e} \mu_e = \eta_e$; and the Hessian matrix, $\partial_{\mathbf{R}_I} \partial_{\mathbf{R}_J} E = \partial_{\mathbf{R}_J} \mathbf{F}_I = \kappa_{IJ}$; the ‘‘force Fukui’’ function $\delta_{Z(\mathbf{r})} \partial_{\mathbf{R}_I} E = -\delta_{Z(\mathbf{r})} \mathbf{F}_I = \partial_{\mathbf{R}_I} \mu_n(\mathbf{r}) = ff(\mathbf{R}_I, \mathbf{r})$; the molecular Fukui function $\partial_{N_e} \delta_{Z(\mathbf{r})} E = \partial_{N_e} \mu_n(\mathbf{r}) = \delta_{Z(\mathbf{r})} \mu_e = f_m(\mathbf{r})$; and the nuclear Fukui function $\partial_{N_e} \partial_{\mathbf{R}_I} E = -\partial_{N_e} \mathbf{F}_I = \partial_{\mathbf{R}_I} \mu_e = f_n(\mathbf{R}_I)$.

The (arbitrary) path along λ can be chosen to simplify the number of terms in Eq. (2). For the small molecules in the first part of this study, we have chosen to relax the geometry of the molecule along λ so that $E(\lambda)$ is minimal with respect to geometrical variation, and the term containing \mathbf{F} vanishes. Furthermore, we restrict ourselves to alchemical transforma-

tions with a constant N_e equal to 8 valence electrons (isoelectronic). For such paths Eq. (2) reduces to

$$\partial_\lambda E = \int d\mathbf{r} \mu_n(\mathbf{r}) \partial_\lambda Z(\mathbf{r}). \quad (3)$$

The alchemical derivative of Eq. (3) describes the change in energy along the path λ as the integral over the change in proton density, $Z(\mathbf{r})$, times the nuclear chemical potential, $\mu_n(\mathbf{r})$.

B. Alchemical derivatives

The nuclear chemical potential, $\mu_n(\mathbf{r})$, was derived in Ref. 10 as an electrostatic potential,

$$\mu_n(\mathbf{r}) = \int d\mathbf{r}' \frac{Z(\mathbf{r}') \text{erf}(\sigma|\mathbf{r} - \mathbf{r}'|) - \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (4)$$

modified with an error function to switch off intra-nuclear interactions. This switching function is required for any \mathbf{r} for which $Z(\mathbf{r}) > 0$, i.e. at an atomic center \mathbf{R}_I . At this point, proton density would be added to an existing nucleus, diverging the electrostatic interaction. In principle, choosing a small value of σ will eliminate such divergences from the integral. In practice, the nuclear charge at \mathbf{R}_I is simply excluded from the evaluation of $\mu_n(\mathbf{R}_I)$.

Changing the proton density at the nuclei is of particular interest because this changes an atom from one element to another. The corresponding molecular nuclear alchemical derivative can be defined as the sum over the product of alchemical potentials with changes in atomic number,

$$\partial_\lambda E = \sum_I \mu_n(\mathbf{R}_I) \partial_\lambda N_I. \quad (5)$$

The focus of this work is to calculate the alchemical derivative of an energy difference, $\Delta E = E^b - E^a$, between two molecular geometries $\{\mathbf{R}_I^b\}$ and $\{\mathbf{R}_I^a\}$. This can be written in terms of a difference in the nuclear chemical potential,

$$\partial_\lambda \Delta E = \int d\mathbf{r} \Delta \mu_n(\mathbf{r}) \partial_\lambda Z(\mathbf{r}), \quad (6)$$

where $\Delta \mu_n(\mathbf{r}) = \mu_n^b(\mathbf{r}) - \mu_n^a(\mathbf{r})$. In the special case that all nuclei have the same position in molecules a and b , so $\{\mathbf{R}_I^b\} = \{\mathbf{R}_I^a\} = \{\mathbf{R}_I\}$, the alchemical potential difference reduces according to Eq. 4 to

$$\begin{aligned} \Delta \mu_n(\mathbf{R}_I) &= \mu_n^b(\mathbf{R}_I) - \mu_n^a(\mathbf{R}_I) \\ &= - \int d\mathbf{r} \frac{\rho^b(\mathbf{r}) - \rho^a(\mathbf{r})}{|\mathbf{R}_I - \mathbf{r}|} \end{aligned} \quad (7)$$

where $\rho^a(\mathbf{r})$ and $\rho^b(\mathbf{r})$ are the electronic charge densities in molecules a and b respectively.

C. Varying molecular geometries

Here we are interested in energy differences between molecules with different geometries, where $\{\mathbf{R}_I^a\} \neq \{\mathbf{R}_I^b\}$.

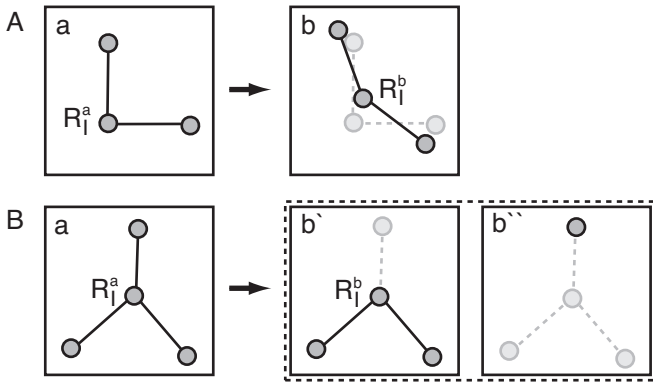


FIG. 1: (A) Activation energies are calculated as the energy difference between two geometries; the corresponding alchemical derivative at atom I is defined in Eq. 9. (B) In a dissociation reaction where atom I is only present in the reactant and product space a and b' , respectively, space b'' does not contribute to the alchemical derivative $\Delta\mu_{n,I}$.

When the position of nucleus I changes, the alchemical potential difference can no longer be defined at the position of the nucleus as $\Delta\mu_n(\mathbf{R}_I)$. Instead, we use the index of the nucleus, I , to identify the alchemical potential difference,

$$\Delta\mu_{n,I} = \mu_n^b(\mathbf{R}_I^b) - \mu_n^a(\mathbf{R}_I^a), \quad (8)$$

and the derivative of the energy difference along λ is

$$\partial_\lambda \Delta E = \sum_I \Delta\mu_{n,I} \partial_\lambda N_I. \quad (9)$$

D. Activation and protonation energies

Two examples of reaction energetics involving molecules of different geometries are illustrated schematically in Fig. 1, an activation energy (A) and a protonation energy (B). The activation energy is the energy difference between two geometries, the initial state $\{\mathbf{R}_I^{\text{init}}\}$ and the transition state $\{\mathbf{R}_I^{\text{ts}}\}$, $E^{\text{act}} = E^{\text{ts}} - E^{\text{init}}$. Its derivative along the path λ is

$$\partial_\lambda E^{\text{act}} = \sum_I \mu_{n,I}^{\text{act}} \partial_\lambda N_I, \quad (10)$$

where $\mu_{n,I}^{\text{act}} = \mu_n^{\text{ts}}(\mathbf{R}_I^{\text{ts}}) - \mu_n^{\text{init}}(\mathbf{R}_I^{\text{init}})$.

The reaction energy E^{rxn} and its alchemical potential at atom I , $\mu_{n,I}^{\text{rxn}}$, are defined in analogy, substituting the final state for the transition state. In this work, we chose the protonation of a molecule as the reaction energy to be calculated,

$$E^{\text{prot}} = E^{\text{mol}+\text{H}^+} - E^{\text{mol}} - E^{\text{H}^+}. \quad (11)$$

This is illustrated in Fig. 1(B) where a is the protonated molecule and b contains both the isolated molecule (b') and proton (b''). When the molecular geometry is relaxed, the alchemical derivatives at the atoms can be calculated using Eq. (8) and the derivative of E^{prot} along λ using Eq. (9). If the

geometry of the molecule is held fixed, the alchemical potential is defined everywhere in space

$$\begin{aligned} \mu_n^{\text{prot}}(\mathbf{r}) &= \delta_{Z(\mathbf{r})} E^{\text{prot}} \\ &= \mu_n^{\text{mol}+\text{H}^+}(\mathbf{r}) - \mu_n^{\text{mol}}(\mathbf{r}) - \mu_n^{\text{H}^+}(\mathbf{r}) \end{aligned} \quad (12)$$

Note that this expression describes the energy of adding proton density anywhere in space, not just at existing atoms. There is, however, a subtlety in Eq. (12) related to how reaction energies are calculated with DFT. The three terms in Eq. (11) are calculated separately as illustrated in Fig. 1(B). Even though the atoms are fixed in space, $\mu_n^{\text{mol}}(\mathbf{r})$ and $\mu_n^{\text{H}^+}(\mathbf{r})$ are calculated in the difference spaces b' and b'' ; they do not share the same coordinate, \mathbf{r} . To calculate an alchemical change in the energy one must decide in which space the change occurs, \mathbf{r}^{mol} or \mathbf{r}^{H^+} . If the change is in \mathbf{r}^{mol} then $\mu_n^{\text{H}^+}$ vanishes in Eq. (12).

E. Alchemical derivatives via finite difference

Eq. (5) can be written as a finite difference derivative,

$$\frac{\Delta E}{\Delta \lambda} = \sum_I \frac{\Delta E}{\Delta N_I} \frac{\Delta N_I}{\Delta \lambda}. \quad (13)$$

Note here that ΔE is a change in energy with respect to λ and not a change with respect to geometry, as in Eq. (6). In standard electronic structure calculations the atomic numbers N_I are integers. The alchemical potential, $\Delta E / \Delta N_I$ could be evaluated with integer N_I , but we are interested in using a smaller (non-integer) finite difference step size in order to numerically verify the analytic derivative from Eq. (4). To this end we choose to represent an atom with a fractional atomic number by a pseudopotential that is linearly interpolated between the atoms of nearest integer atomic numbers. For example, an atom which is fractionally higher by an amount f than the atomic species with integer atomic number N , the pseudopotential is interpolated as

$$V_{N+f}^{PP} = (1-f)V_N^{PP} + fV_{N+1}^{PP}. \quad (14)$$

With interpolated pseudopotentials we can calculate both E and $\mu_{n,I}$ at any non-integer value of N_I and thus the finite difference derivatives for any value of λ along an alchemical transition. Energies were calculated along paths, $\lambda \in [0, 1]$, in increments of 0.1. Finite difference alchemical derivatives were taken at these points using central difference with a step size of $\Delta \lambda = 0.005$.

III. COMPUTATIONAL DETAILS

DFT [23, 26] calculations were preformed using the generalized gradient approximated exchange-correlation potential PBE [27], as implemented in Vienna *ab initio* simulation package, VASP [28, 29]. Core electrons were represented using the projector augmented wave method [30, 31]; valence

electrons with a plane wave basis set up to an energy cut-off of 400 eV. A cubic cell of dimension 10 Å was used for molecules and 20 Å for nanoparticles. Geometry relaxation was continued until the force dropped below 0.01 eV/Å for each atom. In VASP charged cells are with a uniform background counter-charge.

Analytical alchemical potentials were evaluated from Eq. (4) on a uniform grid with 20-30 points/Å using the electrostatic potential output. To reduce finite grid-based errors when taking differences such as in Eq. (8) the position of the atoms \mathbf{R}_I^a and \mathbf{R}_I^b were translated to the same position on the grid.

A calculation of $\partial_\lambda E$ requires a path of alchemical transformation $N_I(\lambda)$ connecting the initial and final atomic number N_I^a and N_I^b . We have chosen paths in which the N_I vary linearly in λ ,

$$N_I(\lambda) = (1 - \lambda) N_I^a + \lambda N_I^b. \quad (15)$$

This choice is convenient because the derivative is a constant along the path,

$$\partial_\lambda N_I(\lambda) = \partial_\lambda N_I = N_I^b - N_I^a, \quad (16)$$

and depends only on the endpoints.

IV. RESULTS AND DISCUSSION

Two test cases are used to verify that the analytic alchemical derivatives are consistent with numerical finite difference derivatives of the pseudopotentials. First, the umbrella flipping of ammonia is used to test the change in activation energies. Then, second, the protonation of small molecules is used to test derivatives in reaction enthalpies for both frozen and relaxed geometries. In both cases, there is good agreement between the analytical and numerical alchemical derivatives. Finally, we show a more interesting example of oxygen binding to a palladium nanoparticle. This final case is motivated by Nørskov and coworkers who showed that the binding energy of oxygen to a metal surface correlates to its catalytic activity for the oxygen reduction reaction [19]. Here, we test the predictive power of the alchemical derivatives for integer changes in the atomic number of the metal atoms in the nanoparticle, and show that the derivatives can be used to tune the binding energy of oxygen with respect to the particle composition.

A. Activation energies

The energy barrier is a geometry dependent property where the configurations of the transition state and initial state differ and as such explicitly requires the nuclei mapping described in Eq. 8 to calculate the alchemical potential. We have calculated the barrier for the umbrella flipping of ammonia. NH_3 has two degenerate energy minima, a pyramidal structure with C_{3h} symmetry and its mirror plane conformer. Figure 2(A) shows a nudged elastic band [32, 33] calculation of the flipping barrier with a planar transition state and an activation energy of 0.21 eV.

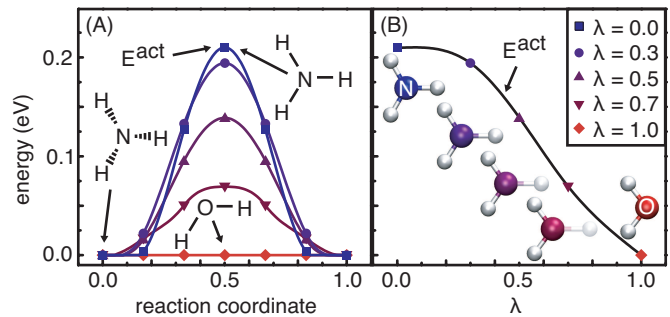


FIG. 2: (A) The minimum energy path for the umbrella inversion of NH_3 passes through a planar transition state. This path is shown for several values of λ , where NH_3 ($\lambda = 0$) is alchemically transformed to H_2O ($\lambda = 1$). (B) E^{act} for the inversion process as a function of λ . The structures of the transition states are shown for select values of λ .

We are interested in how the flipping barrier changes as we alchemically transform NH_3 into a different molecule – H_2O in this example. Water was chosen in part because it is a neighbor to ammonia in the alchemical space of molecules with unit changes in atomic numbers, and also because the flipping process for H_2O becomes a rotation with no energy barrier. This provides a convenient test case for the calculation of activation energies and the corresponding derivatives along an alchemical transformation where in the limit of $\lambda \rightarrow 1$, E^{act} is known *a priori* to be zero.

The alchemical transformation is defined by λ going from 0 to 1 as NH_3 is mutated to H_2O . We have chosen the transformation to correspond to the simultaneous addition of a proton the central atom and the annihilation of a H nucleus, linearly, as defined in Eq. (15). Figure 2(B) shows how the activation energy for the inversion barrier of NH_3 first decreases only slightly, then decays more rapidly, and finally vanishes as the molecule is transformed into H_2O .

The usefulness of alchemical derivatives for compound design can be measured by how well they can extrapolate to integer changes in N_I , corresponding to elements which exist in nature. For the alchemical transformation between NH_3 and H_2O , the $\{N_I\}$ change linearly in λ [Eq. (15)], but the property, E^{act} has significant curvature along λ . In Fig. 2(B), one can see how the linear extrapolation based on the derivative of E^{act} with respect to λ at H_2O would yield a reasonable prediction for the barrier of NH_3 , whereas the derivative at NH_3 would significantly overestimate the flipping barrier at H_2O . Current efforts aim to determine alchemical paths that have first order derivatives that are more amenable to extrapolations to new compounds [12].

In previous work, relaxation of the force $\{\mathbf{F}_I\}$ for points along λ was only done for one structure because of the constraint, $\{\mathbf{R}_I^a\} = \{\mathbf{R}_I^b\}$ [34]. Using Eq. (8), this constraint is lifted, and the structures are relaxed independently. The most prominent geometry change of both the transition state and the initial state as ammonia is converted to water along λ , is the bond-length between the central atom and the hydrogen being annihilated. Figure 3 shows that the bond-length increases at

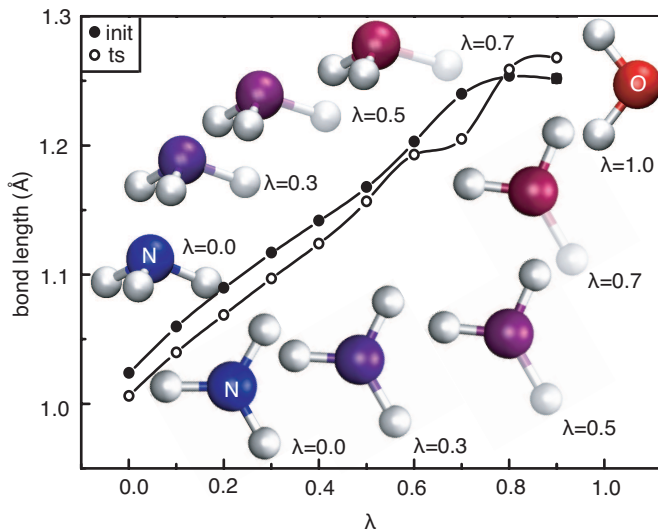


FIG. 3: The bond distance between the central atom and the hydrogen being annihilated increases along λ for both the relaxed initial structure $\{\mathbf{R}_i^{\text{init}}\}$ and the relaxed transition state $\{\mathbf{R}_i^{\text{ts}}\}$ for the umbrella flipping process.

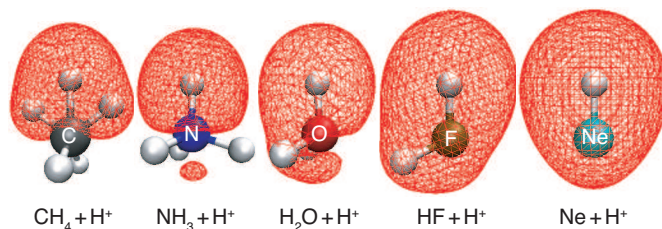


FIG. 4: Isosurfaces of the alchemical potential $\mu_n^{\text{prot}}(\mathbf{r}) = 7$ eV from Eq. (12) are shown for the protonated methane series using geometries fixed to NH_4^+ .

the same rate for both the optimized initial state and the converged transition state. As the bond lengthens, the interaction between the fractional proton and the rest of the molecule decreases until it vanishes at $\lambda = 1$.

B. Protonation energies

The second geometry-dependent property we discuss is the protonation energy for a set of iso-electronic molecules, CH_4 , NH_3 , H_2O , HF , and Ne . In the case of CH_4 , for example, the protonation energy defined in Eq. (11) is $E^{\text{prot}} = E^{\text{CH}_4 + \text{H}^+} - E^{\text{CH}_4} - E^{\text{H}^+}$. The choice of this property is motivated by the fact that it can be evaluated both for frozen and relaxed geometries, allowing to measure the effect of geometry relaxation when calculating alchemical derivatives.

In the case of fixed geometries we chose the relaxed NH_4^+ structure as the common structure for all of the molecules. For different molecules, such as H_2O , the central atom, O, was located at the N position and the two hydrogen atoms were in two of the four equivalent H positions. In the case of CH_5^+ an extra hydrogen atom was relaxed in the fixed NH_4^+ structure, and then kept fixed for its alchemical transformation.

With fixed geometries the alchemical potential for protonation is defined over all space according to Eq. (12). This is a useful function because it shows where a new atom should be added to change a desired property—in this case the protonation energy. Figure 4 shows isosurfaces of the alchemical potential $\mu_n^{\text{prot}}(\mathbf{r})$ for all the five compounds. Inside each of these surfaces, the change in energy is over 7 eV/proton for the molecule shown as compared to the de-protonated molecule without the uppermost proton. We note some features of the alchemical potential. First, the potential is very high around the uppermost proton because this is the proton which was added to calculate the protonation energy. Adding further proton density near this same location gives rise to the high electrostatic potential. Second, as one would expect from the Coulomb-explosion of water, the alchemical potential is always positive, indicating that at no point in space does the alchemical addition of an extra proton to the system favor the protonation energy—the energy of protonation increases as protons are added.

The alchemical change we have considered is an increase in the nuclear number of the central atom. At the same time, the proton in a bonded H atom is eliminated so the total number of protons is constant over the sequence of alchemical transformations, $\text{CH}_4 \rightarrow \text{NH}_3$, $\text{NH}_3 \rightarrow \text{H}_2\text{O}$, $\text{H}_2\text{O} \rightarrow \text{HF}$, and $\text{HF} \rightarrow \text{Ne}$. The alchemical potential of protonation is positive for each of these molecules, but the alchemical derivative along these paths also includes the proton annihilation. The change in protonation energy along these paths is the difference in alchemical potentials at the central atom and at the proton being removed. Figure 5 shows how the protonation energies change along the alchemical transformations, and how the difference in energy between the unfavorable addition of a proton at the center nucleus and the favorable H annihilation can lead to a lower overall energy. Fractional nuclei are treated using the numerical interpolations described in Sec. II E so that energies can be calculated along the alchemical mutation.

2. Relaxed geometries

The effect of molecular relaxation is shown in Fig. 5 to be particularly significant for the first transformation between CH_4 and NH_3 . There are two possible alchemical transformations between these molecules in a fixed geometry case because the protonated CH_5^+ molecule has symmetrically nonequivalent protons, those which are neighboring the protonating H^+ and those which are not. The two different alchemical paths correspond to which type of proton is being

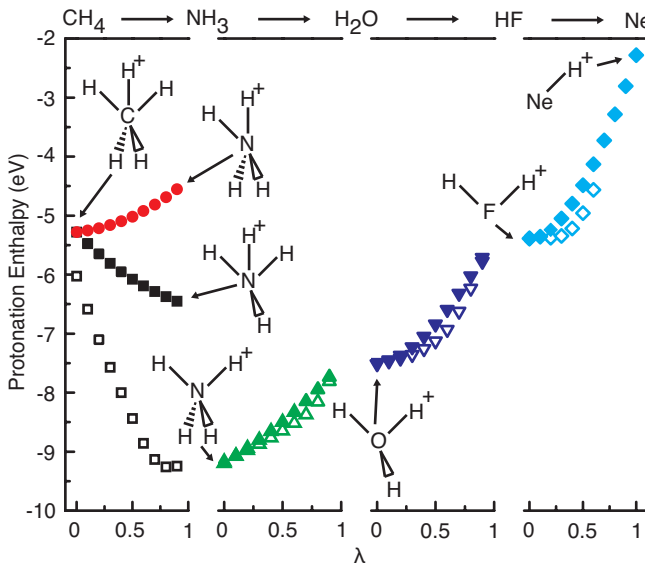


FIG. 5: Left to right is the protonation energy as CH_4 is alchemically transformed into Ne. The structures shown as insets are the protonated endpoints whose pseudopotentials have been interpolated when following alchemical paths. For each alchemical path the number of protons on the central nucleus is increased by one and a hydrogen nucleus is concurrently annihilated. The solid symbols represent configurations fixed to the optimal NH_4^+ geometry and the empty symbols are relaxed. A consequence of using fixed structures is that all hydrogens are symmetrically equivalent for the CH_5^+ geometry; two separate paths emerge by selecting which hydrogen is to be annihilated.

annihilated. Upon relaxation, the asymmetry is removed as the molecules move to a single minimum energy structure. As one would expect, the difference between the fixed-geometry and relaxed-geometry paths are most significant for molecules which are highly strained, such as the NH_4^+ endpoint of the first transition. In the subsequent transitions, the fixed and relaxed geometries are much closer. As in the case of ammonia inversion, there is a significant curvature in the alchemical paths which reduces the predictive quality of the alchemical derivatives for integer changes in N_I .

C. Alchemical potentials

To separate and validate the derivatives in Eq. (2) we focus only on the alchemical term, $\partial_\lambda E = \int d\mathbf{r} \mu_n(\mathbf{r}) \partial_\lambda Z(\mathbf{r})$, as specified in Eq. (3). This is accomplished by choosing a system where N_e is constant, i.e. $\partial_\lambda N_e = 0$. The second term in Eq. (2) is also zeroed by relaxing the ions along the path so $\{\mathbf{F}_I\} = 0$. We have a particular interest in validating the analytical calculation of $\Delta\mu_{n,I}$ from Eq. (7) because it is from this calculation that $\partial_\lambda \Delta E$ depends, to first order.

In this study, every transforming path involves the alchemical mutation of two atoms, the simultaneous annihilation of a proton in a hydrogen atom and the increase of the atomic number of a central atom to which it is bound. As such, Eq. (9) is only non-zero for changes at the central atom and the anni-

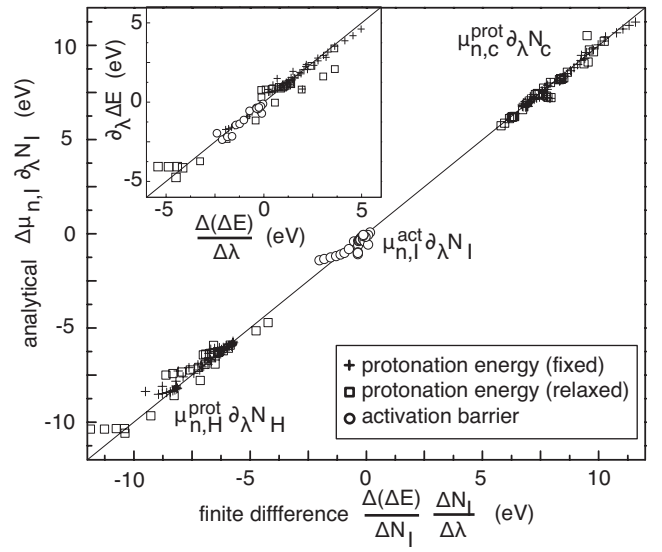


FIG. 6: Correlation of the analytical [Eq. (9)] and numerical [Eq. (13)] alchemical derivatives $\Delta\mu_{n,H} \partial_\lambda N_H$ and $\Delta\mu_{n,c} \partial_\lambda N_c$, for all paths and all values of λ . The inset shows a similar correlation for the sum of these atomic components, which give derivatives of the binding and activation energies along the alchemical paths in Table I.

lated hydrogen and reduces to $\partial_\lambda E = \Delta\mu_{n,H} \partial_\lambda N_H + \Delta\mu_{n,c} \partial_\lambda N_c$, where H and c indicate the hydrogen and central atom, respectively. For our linear transformations, $\partial_\lambda N_c = -\partial_\lambda N_H = 1$.

In Fig. 6 the accuracy of the individual terms $\Delta\mu_{n,H} \partial_\lambda N_H$ and $\Delta\mu_{n,c} \partial_\lambda N_c$ are compared separately to finite difference derivatives in λ as obtained from Eq. (13). The data is for all paths at each intermediate λ values considered for this study. For the two applications, barrier and reaction enthalpy, we have computed the alchemical potential terms according to Eqs. (10, 12). We find a correlation with a root-mean-square deviation of 0.29 eV for the barriers, and 0.26 eV (fixed-geometry) and 0.44 eV (relaxed-geometry) for the reaction enthalpies. The remaining deviation from the finite difference results is most likely due to the pseudopotential interpolation used to calculate $\frac{\Delta N_I}{\Delta \lambda}$ which is not the same as the analytical $\partial_\lambda Z$. Nonetheless, the pseudopotential is designed to model an atom with atomic number Z , so that linear changes in the pseudopotential are a reasonable approximation to linear changes in Z .

D. Oxygen binding on a Pd nanoparticle

For the alchemical potential to be useful for rational compound design it must be predictive for changes in the property of interest between real materials. Consequently, the potential must be sufficiently accurate in predicting changes to the property, and ideally the property should depend linearly on N_I . The numerical accuracy of the analytical alchemical potential has been demonstrated in the previous sections. Here, we address the predictive quality for the binding of molecular oxygen to the hollow site in the center of the (111) face of a

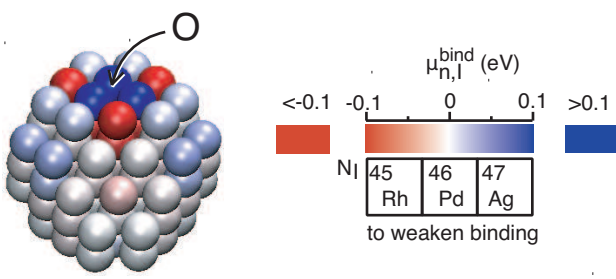


FIG. 7: The alchemical potential $\mu_{n,l}^{\text{bind}}$, calculated from Eqs. (8,17), for the binding of O to a 79 atom Pd nanoparticle. Colors represent the value of the alchemical potential on each atom. To weaken oxygen binding atoms colored in red should be changed to the more noble Ag and atoms colored in blue should be changed to Rh. The opposite transformations should be made to strengthen the oxygen bond.

79 atom Pd nanoparticle. The property can be expressed in the form of Eq. (11) as

$$E^{\text{bind}} = E^{\text{Pd79+O}} - E^{\text{Pd79}} - \frac{1}{2}E^{\text{O}_2}. \quad (17)$$

As in the previous examples, we isolate the first, nuclear term of Eq. (2) by relaxing all geometries ($\{\mathbf{F}_I\} = 0$), and by performing exclusively iso-electronic changes ($dN_e = 0$). We constrain all transitions to charge neutral paths by pairing the addition of a proton to any nucleus with the removal of a proton from a second nucleus in the system. Furthermore, we restrict ourselves to all geometrically stable and complementary mutant pairs with D_{4h} -symmetry as illustrated in right hand panel of Figure 8.

Figure 7 shows the alchemical potential for the binding energy of oxygen to the nanoparticle. The extreme values for $\mu_{n,l}^{\text{bind}}$ are on the (111) face where the O-atom binds; they fall off with distance from the binding site. If the goal is to tune the binding energy to a desired value, $\mu_{n,l}^{\text{bind}}$ can be used to decide which atoms in the system to alchemically change and in what direction. The best pure-metal for the oxygen reduction reaction is Pt, which binds O more weakly than Pd. To weaken the binding, Fig. 7 shows that the atoms directly at the binding site should be made more noble by increasing N_I . Alternatively, the second-neighbors can be made less noble by decreasing N_I , which is consistent with the ligand effect of near surface alloys [35–37].

Figure 8 shows the correlation between the actual change in calculated binding energy due to mutation, ΔE^{bind} , and the predicted change in binding energy, $\partial_\lambda E^{\text{bind}} d\lambda$ for $d\lambda = 1$, based on the sum over alchemical potentials in Eq. (9). Final structures A and B have 24, structures C-H have 48, and structures I-J have 72, nuclei alchemically changed from the initial 79 atom Pd particle, respectively. All actual changes have been obtained for geometrically relaxed complexes. The correlation between the prediction from the alchemical potential evaluated at the initial Pd particle and the calculated change in O binding for the alchemically modified final particles is very reasonable. Even for changes of up to 72 of the 79 atoms in the Pd particle, the alchemical potentials give a decent prediction

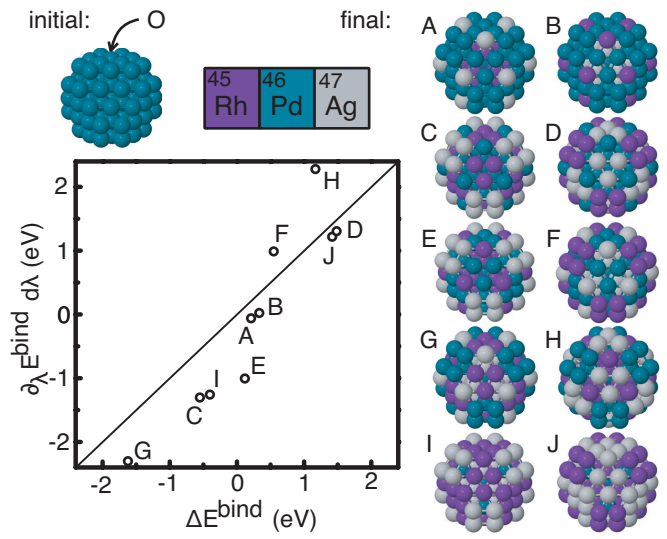


FIG. 8: Correlation of alchemically predicted changes, $\partial_\lambda E^{\text{bind}} d\lambda$ for $d\lambda = 1$, with actual changes to the binding energy, ΔE^{bind} , of molecular oxygen to a 79 atom Pd cluster (initial) due to integer variations in atomic numbers (final). The final systems correspond to complementary mutant pairs with D_{4h} -symmetry.

for the change in the O binding energy. We emphasize that the computational cost for obtaining the predicted changes in binding energy correspond to a single evaluation of the binding energy for oxygen to the initial 79 atom Pd cluster. Albeit higher order terms in $\{Z_I\}$ and $\{R_I\}$ of the Taylor expansion in Eq. (1) could be used to improve the correlation, their associated computational cost to evaluate the response density, e.g. within linear response theory [38, 39], defeats the original purpose of leveraging exclusively analytical gradients altogether.

V. CONCLUSIONS

We have presented numerical evidence that accurate analytical derivatives can be computed for chemical properties that involve variations in geometry. While test systems show some non-linearity in the alchemical paths, the derivatives can be predictive for integer changes in the number of protons as seen in the binding energy of oxygen to a Pd nanoparticle. Furthermore, once the the property of interest has been calculated for the initial compound, the corresponding alchemical derivative can be evaluated with negligible additional cost. Then, the effect of alchemical changes throughout the system can be estimated without doing additional calculations for all the various possible final compounds. We note that all the alchemical derivatives required for the results that we report and discuss in Section IV D are in the standard output of the employed computer code VASP [28, 29]. This is a promising initial step towards the use of alchemical derivatives for the gradient-based optimization of materials.

On a more fundamental level, this study has shown how to deal at the electronic structure level with geometrically de-

pendent properties of systems that are subjected to alchemical changes. Consequently, it appears also possible to calculate the predicted effects of alchemical changes on finite temperature dependent properties, such as free energies of binding or barriers, with insignificant computational overhead, using statistical mechanical sampling methods of phase space, e.g. *ab initio* molecular dynamics [40].

A. Acknowledgments

This work was funded by the National Science Foundation (CHE-0645497), the Department of Energy under con-

tract DE-FG02-09ER16090 and the Welch Foundation (F-1601). D. S. is grateful for support from the Sandia National Laboratory (SNL) summer student internship program at the Computer Science Research Institute. OAvL acknowledges support from SNL Truman Program Laboratory Directed Research Development project No. 120209. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. We gratefully acknowledge the Texas Advanced Computing Center for computational resources.

-
- [1] L. C. Blum and J.-L. Reymond, *J. Am. Chem. Soc.* **131**, 8732 (2009).
 - [2] T. Fink and J.-L. Reymond, *J. Chem. Inf. Model.* **47**, 342 (2007).
 - [3] B. K. Shoichet, *Nature* **432**, 862 (2004).
 - [4] R. Guha, D. Dutta, P. Jurs, and T. Chen, *jcim* **46**, 1836 (2006).
 - [5] S. Zhang, A. Golbraikh, S. Oloff, H. Kohn, and A. Tropsha, *J. Chem. Inf. Model.* **46**, 1984 (2006).
 - [6] A. Franceschetti and A. Zunger, *Nature* **402**, 60 (1999).
 - [7] G. H. Jönhannesson, T. Bligaard, A. V. Ruban, H. L. Skriver, K. W. Jacobsen, and J. K. Nørskov, *Phys. Rev. Lett.* **88**, 255506 (2002).
 - [8] T. Mueller and G. Ceder, *Phys. Rev. B* **80**, 024103 (2009).
 - [9] O. A. von Lilienfeld, R. Lins, and U. Rothlisberger, *Phys. Rev. Lett.* **95**, 153002 (2005).
 - [10] O. A. von Lilienfeld and M. E. Tuckerman, *J. Chem. Phys.* **125**, 154104 (2006).
 - [11] V. Marcon, O. A. von Lilienfeld, and D. Andrienko, *J. Chem. Phys.* **127**, 064305 (2007).
 - [12] O. A. von Lilienfeld, *J. Chem. Phys.* **131**, 164102 (2009).
 - [13] M. Wang, X. Hu, D. N. Beratan, and W. Yang, *J. Am. Chem. Soc.* **128**, 3228 (2006).
 - [14] D. Xiao, W. Yang, and D. N. Beratan, *J. Chem. Phys.* **129**, 044106 (2008).
 - [15] X. Hu, D. N. Beratan, and W. Yang, *J. Chem. Phys.* **129**, 064102 (2008).
 - [16] D. Balamurugan, W. Yang, and D. N. Beratan, *J. Chem. Phys.* **129**, 174105 (2008).
 - [17] S. Keinan, M. J. Therien, D. N. Beratan, and W. Yang, *J. Phys. Chem. A* **112**, 12203 (2008).
 - [18] B. C. Rinderspacher, J. Andzelm, A. Rawlett, J. Dougherty, D. N. Beratan, and W. Yang, *J. Chem. Theory Comput.* **5**, 3321 (2009).
 - [19] J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard, and H. Jónsson, *J. Phys. Chem. B* **108**, 17886 (2004).
 - [20] J. G. Kirkwood, *J. Chem. Phys.* **3**, 300 (1935).
 - [21] R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
 - [22] W. F. van Gunsteren, X. Daurab, and A. E. Markc, *Helv. Chim. Acta* **85**, 3113 (2002).
 - [23] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
 - [24] J. F. Janak, *Phys. Rev. B* **18**, 7165 (1978).
 - [25] P. Mori-Sánchez, A. J. Cohen, and W. Yang, *Phys. Rev. Lett.* **102**, 066403 (2009).
 - [26] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
 - [27] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
 - [28] G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
 - [29] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
 - [30] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
 - [31] G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
 - [32] G. Henkelman, B. P. Uberuaga, and H. Jónsson, *J. Chem. Phys.* **113**, 9901 (2000).
 - [33] D. Sheppard, R. Terrell, and G. Henkelman, *J. Chem. Phys.* **128**, 123106 (2008).
 - [34] O. A. von Lilienfeld and M. E. Tuckerman, *J. Chem. Theory Comput.* **3**, 1083 (2007).
 - [35] J. Greeley, J. K. Nørskov, and M. Mavrikakis, *Annu. Rev. Phys. Chem.* **53**, 319 (2002).
 - [36] J. Zhang, M. B. Vukmirovic, K. Sasaki, A. U. Nilekar, M. Mavrikakis, and R. R. Adzic, *J. Am. Chem. Soc.* **127**, 12480 (2005).
 - [37] W. Tang and G. Henkelman, *J. Chem. Phys.* **130**, 194505 (2009).
 - [38] A. Putrino, D. Sebastiani, and M. Parrinello, *J. Chem. Phys.* **113**, 7102 (2000).
 - [39] O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, *J. Chem. Phys.* **122**, 014113 (2005).
 - [40] R. Iftimie, P. Minari, and M. E. Tuckerman, *Proc. Natl. Acad. Sci. USA* **102**, 6654 (2005).